

# A VERY ELEMENTARY PROOF OF THE B. AND M. SHAPIRO CONJECTURE FOR CUBIC RATIONAL FUNCTIONS

XANDER FABER AND BIANCA THOMPSON

**ABSTRACT.** Using essentially only algebra, we give a proof that a cubic rational function over  $\mathbb{C}$  with real critical points is equivalent to a real rational function. We also show that the natural generalization to  $\mathbb{Q}_p$  fails unless  $p = 3$ .

## 1. INTRODUCTION

Let  $K$  be a field of characteristic zero with algebraic closure  $\bar{K}$ . We say that two rational functions  $f, g \in \bar{K}(z)$  are **equivalent** if there is a fractional linear transformation  $\sigma \in \bar{K}(z)$  such that  $f = \sigma \circ g$ . Viewing  $f$  and  $g$  as endomorphisms of the projective line, we see that they are equivalent if they differ by a change of coordinate on the target. Note that equivalent rational functions have the same critical points.

**Theorem 1.1** (Eremenko/Gabrielov). *If  $f \in \mathbb{C}(z)$  is a rational function with real critical points, then  $f$  is equivalent to a rational function with real coefficients.*

By relating equivalence classes of rational functions with special Schubert cycles, Goldberg [3] showed that there are at most  $\rho(d) := \frac{1}{d} \binom{2d-2}{d-1}$  equivalence classes of degree  $d$  rational functions with a given set of critical points. Eremenko and Gabrielov [1, 2] used topological, combinatorial, and complex analytic techniques to *construct* exactly  $\rho(d)$  real rational functions with a given set of real critical points, which proves the theorem.

But the correspondence between a rational function and its critical points is purely algebraic, via roots of the derivative. This raises the question of whether a truly elementary proof of the above result exists — one that does not use any analysis or topology. We give such a proof for cubic functions in this note.

*Remark 1.2.* The quadratic case is trivial over any field: direct computation shows that a function with critical points  $c_1, c_2 \in \mathbb{P}^1(K)$ ,  $c_1 \neq \infty$ , is equivalent to either  $(z - c_1)^2$  or  $\left(\frac{z-c_1}{z-c_2}\right)^2$ , depending on whether or not  $c_2 = \infty$ .

For a field  $K$  and a nonconstant rational function  $\phi \in K(z)$ , we say that  $K$  is  $\phi$ -**perfect** if the map  $\phi: \mathbb{P}^1(K) \rightarrow \mathbb{P}^1(K)$  is surjective. For example, if  $K$  has characteristic  $p$  and  $\phi(z) = z^p$ , then  $K$  is  $\phi$ -perfect if and only if it is a perfect field in the usual sense.

**Theorem 1.3.** *Let  $K$  be a field of characteristic zero with algebraic closure  $\bar{K}$ . The following statements are equivalent:*

- (1) *Any cubic rational function  $f \in \bar{K}(z)$  with  $K$ -rational critical points is equivalent to a rational function in  $K(z)$ .*
- (2)  *$K$  is  $\phi$ -perfect, where  $\phi(z) = -\frac{z^2+2z}{2z+3}$ .*

The theorem will be proved in §2.

---

*Key words and phrases.* critical points, cubic rational functions, B. and M. Shapiro conjecture.

**Corollary 1.4.** *If  $f \in \mathbb{C}(z)$  is a cubic rational function with real critical points, then  $f$  is equivalent to a real rational function.*

*Proof.* The field of real numbers is  $\phi$ -perfect for  $\phi$  as in the theorem. Indeed,  $\phi(-2/3) = \infty$ , and if  $y \in \mathbb{R}$ , then the equation  $\phi(z) = y$  is equivalent to a quadratic equation with discriminant  $4(y^2 - y + 1) = (2y - 1)^2 + 3 > 0$ . Hence  $\phi(z) = y$  has a real solution.  $\square$

For what other fields of interest  $K$  does the corollary on equivalence of rational functions continue to hold? Said another way, which fields  $K$  are  $\phi$ -perfect for  $\phi(z) = -\frac{z^2+2z}{2z+3}$ ?

For one source of (non-)examples, we look at non-Archimedean completions of the rational numbers.

**Proposition 1.5.** *Set  $\phi(z) = -\frac{z^2+2z}{2z+3}$ . The field  $\mathbb{Q}_p$  is  $\phi$ -perfect if and only if  $p = 3$ .*

This proposition will be proved in §3. Using Theorem 1.3, we obtain the following as a consequence:

**Corollary 1.6.** *Let  $p$  be a prime.*

- *If  $p \neq 3$ , then there exists a cubic rational function  $f \in \bar{\mathbb{Q}}_p(z)$  that has  $\mathbb{Q}_p$ -rational critical points, but that is not equivalent to a rational function with coefficients in  $\mathbb{Q}_p$ .*
- *If  $f \in \bar{\mathbb{Q}}_3(z)$  is a cubic rational function with  $\mathbb{Q}_3$ -rational critical points, then  $f$  is equivalent to a rational function with coefficients in  $\mathbb{Q}_3$ .*

## 2. PROOF OF THE THEOREM

We begin with a normal form for cubic functions. For  $u \in \bar{K} \setminus \{-1, -2\}$ , define

$$f_u(z) = \frac{z^2(z+u)}{(2u+3)z - (u+2)}. \quad (2.1)$$

(We exclude  $u = -1, -2$  because otherwise a root of the numerator and the denominator collide, and  $f_u$  degenerates to a quadratic.) This function has the property that it fixes 0, 1, and  $\infty$ , and each of these three points is critical.

**Lemma 2.1.** *A cubic rational function that is critical at 0, 1, and  $\infty$  is equivalent to a unique  $f_u$ , and the fourth critical point is  $\phi(u) = -\frac{u^2+2u}{2u+3}$ .*

*Proof.* Write  $f$  for a cubic function that is critical at 0, 1, and  $\infty$ . By a change of coordinate on the target, we may assume that 0, 1, and  $\infty$  are all fixed points, in which case  $f$  is of the form

$$f(z) = \frac{z^3 + uz^2}{vz + (u - v + 1)}$$

for some  $u, v \in \bar{K}$ . The Wronskian has the form

$$z(2vz^2 + (uv + 3(u - v + 1))z + 2u(u - v + 1)).$$

Substituting  $z = 1$  kills this expression (since 1 is a critical point). Solving the resulting equation for  $v$  yields  $v = 2u + 3$ . Hence  $f$  is equivalent to (2.1), as desired.

For the uniqueness statement, suppose that  $f_u$  is equivalent to  $f_v$  for some  $u, v$ . Then there is a fractional linear  $\sigma \in \bar{K}(z)$  such that  $f_u = \sigma \circ f_v$ . But  $f_u$  and  $f_v$  both fix 0, 1, and  $\infty$ , so that  $\sigma$  does as well. This means  $\sigma(z) = z$ , and  $u = v$ .

The fourth critical point of  $f_u$  may be found by factoring the derivative.  $\square$

*Remark 2.2.* Note that taking  $u = 0, -3$ , or  $-3/2$  gives a double critical point at 0, 1, or  $\infty$ , respectively.

**Proposition 2.3.** *If  $f_u \in \bar{K}(z)$  is equivalent to a rational function with  $K$ -coefficients, then  $u \in K$ .*

*Proof.* Let  $\sigma \in \bar{K}(z)$  be a fractional linear map such that  $\sigma \circ f_u$  has coefficients in  $K$ . The images of  $0, 1$ , and  $\infty$  under  $\sigma \circ f_u$  all lie in  $\mathbb{P}^1(K)$ . We may therefore apply a further fractional linear transformation  $\tau$  with  $K$ -coefficients so that  $\tau \circ \sigma \circ f_u$  fixes  $0, 1$ , and  $\infty$ . That is,  $\tau \circ \sigma \circ f_u = f_v$  for some  $v$ . Since  $\tau$  and  $\sigma \circ f_u$  have  $K$ -coefficients, we know that  $v \in K$ . By uniqueness in the lemma, we conclude that  $u = v$ .  $\square$

We are now ready for the proof of the theorem. To prove the implication (1)  $\Rightarrow$  (2), we take  $y \in K$  and attempt to solve the equation  $\phi(u) = y$  with  $u \in K$ . If  $y = \infty$ , then we may take  $u = -3/2$ . Otherwise, choose  $u \in \bar{K}$  such that  $\phi(u) = y$ . Then the function  $f_u$  has  $K$ -rational critical points  $\{0, 1, \infty, y\}$ . By (1),  $f_u$  is equivalent to a rational function with  $K$ -coefficients. The above proposition implies that  $u \in K$ .

To prove (2)  $\Rightarrow$  (1), we start with a rational function  $f \in \bar{K}(z)$  with  $K$ -rational critical points. If  $f$  has only two critical points, then each must have multiplicity 2 (by the Riemann-Hurwitz formula). Without loss, we assume they are  $0$  and  $\infty$ , and that  $0, \infty$  are fixed by  $f$ , so that  $f(z) = az^2$  for some  $a \in \bar{K}$ . Evidently  $a^{-1}f$  has coefficients in  $K$ .

Now suppose that  $f$  has at least three critical points. Without loss, we may assume that  $0, 1$ , and  $\infty$  are among them. In particular, by the lemma we see that  $f$  is equivalent to  $f_u$  for some  $u \in \bar{K}$ . The remaining critical point is  $\phi(u) \in K$ . By (2), both solutions of  $\phi(z) = \phi(u)$  lie in  $\mathbb{P}^1(K)$ , so that  $u \in K$ . That is,  $f$  is equivalent to a rational function with  $K$ -coefficients.

### 3. $p$ -ADIC FIELDS

Our proof of Proposition 1.5 is split into the subcases  $p = 2$ ,  $p = 3$ , and  $p > 3$ . Recall that we want to show that  $\mathbb{Q}_p$  is  $\phi$ -perfect for  $\phi(z) = -\frac{z^2+2z}{2z+3}$  if and only if  $p = 3$ . This amounts to determining whether or not  $\phi(z) = y$  does or does not have a solution in  $\mathbb{P}^1(\mathbb{Q}_p)$  for  $y \in \mathbb{Q}_p$ . Rearranging gives the quadratic equation  $z^2 + 2(1+y)z + 3y = 0$ , which has discriminant

$$\Delta = 4(y^2 - y + 1) = (2y - 1)^2 + 3. \quad (3.1)$$

Determining if  $\mathbb{Q}_p$  is  $\phi$ -perfect now amounts to determining whether  $\Delta$  is a square in  $\mathbb{Q}_p$  or not for every  $y \in \mathbb{Q}_p$ .

For  $p = 2$ , set  $y = \frac{1}{2} + t$  with  $t \in \mathbb{Z}_2$ . Then (3.1) becomes

$$\Delta = 4t^2 + 3 \equiv 3 \pmod{4},$$

which is not a square in  $\mathbb{Q}_2$ . Hence  $\phi(z) = \frac{1}{2} + t$  has no solution, and  $\mathbb{Q}_2$  is not  $\phi$ -perfect. (It is worth noting that what we have really proved is that the image of  $\mathbb{P}^1(\mathbb{Q}_2)$  under  $\phi$  is disjoint from the set  $\frac{1}{2} + \mathbb{Z}_2$ .)

For  $p = 3$ , we claim that the discriminant  $\Delta$  is a square in  $\mathbb{Q}_3$  for any  $y \in \mathbb{Q}_3$ . If  $y \in \mathbb{Z}_3$ , then (3.1) shows that  $\Delta$  is a square modulo 3, which is equivalent to  $\Delta$  being a square in  $\mathbb{Z}_3$  by Hensel's lemma. If  $y \notin \mathbb{Z}_3$ , write  $y = y'/3^r$  for some  $y' \in \mathbb{Z}_3^\times$  and  $r > 0$ . Then (3.1) shows that  $\text{ord}_3(\Delta) = -2r$ . Multiplying through by  $3^{2r}$  gives

$$3^{2r}\Delta = (2y3^r - 3^r)^2 + 3^{2r+1} \equiv (2y')^2 \pmod{3}.$$

This last quantity is a nonzero square, from which we deduce that  $3^{2r}\Delta$  is a square in  $\mathbb{Z}_3$  by Hensel's lemma. We conclude that  $\Delta$  is a square in  $\mathbb{Q}_3$ , and  $\mathbb{Q}_3$  is  $\phi$ -perfect.

Finally, we treat the case  $p > 3$ . The resultant of  $\phi(z) = -\frac{z^2+2z}{2z+3}$  is 3, so this rational function may be reduced modulo  $p$  to yield a quadratic function  $\tilde{\phi} \in \mathbb{F}_p(z)$ . Note that  $\tilde{\phi}(0) = \tilde{\phi}(-2)$ , so that  $\tilde{\phi}$  fails to be injective on  $\mathbb{P}^1(\mathbb{F}_p)$ . As  $\mathbb{P}^1(\mathbb{F}_p)$  is a finite set,  $\tilde{\phi}$  also fails to be surjective. Choose  $\tilde{y} \in \mathbb{F}_p$  such that  $\tilde{\phi}(z) = \tilde{y}$  has no solution in  $\mathbb{F}_p$ , and choose a lift  $y \in \mathbb{Z}_p$  such that  $y \equiv \tilde{y} \pmod{p}$ .

It follows that  $\phi(z) = y$  has no solution in  $\mathbb{Z}_p$ . It remains to show that  $\phi(z) = y$  has no solution in  $\mathbb{Q}_p \setminus \mathbb{Z}_p$ . If  $\phi(x) = y$  with  $|x|_p > 1$ , then

$$|\phi(x)|_p = |x|_p \cdot \left| \frac{1 + 2/x}{2 + 3/x} \right|_p = |x|_p > 1,$$

which contradicts  $y \in \mathbb{Z}_p$ . Hence  $\phi(z) = y$  has no solution in  $\mathbb{P}^1(\mathbb{Q}_p)$ , and we have proved that  $\mathbb{Q}_p$  is not  $\phi$ -perfect.

#### 4. FURTHER THOUGHTS

A general rational function of degree  $d > 2$  has  $2d + 1$  free parameters (coefficients) and  $2d - 2$  critical points. Imposing the condition that  $0, 1, \infty$  are fixed and critical reduces this to  $2d - 5$  free parameters. If we fix a set of  $K$ -rational critical points and look at the Wronskian, then the  $2d - 5$  free coefficients for the function must satisfy  $2d - 5$  quadratic equations in  $2d - 5$  variables over  $K$ . In the case  $d = 3$ , in which  $2d - 5 = 1$ , we were able to explicitly solve for the remaining critical point as an explicit function of the free parameter. Is it possible to solve for the critical points as explicit functions of the parameters for  $d > 3$ ?

Bézout's theorem gives an upper bound of  $2^{2d-5}$  solutions for a general system of  $2d - 5$  conics, while Goldberg [3] bounds the number of distinct solutions by the smaller quantity

$$\frac{1}{d} \binom{2d-2}{d-1} \approx \frac{8}{\sqrt{\pi} d^{3/2}} 2^{2d-5}.$$

This suggests a substantial amount of extra structure in our system of equations, which may make it possible to give elementary proofs of the theorem of Eremenko/Gabrielov in degree  $d$  for other small  $d > 3$ .

#### REFERENCES

- [1] A. Eremenko and A. Gabrielov. Rational functions with real critical points and the B. and M. Shapiro conjecture in real enumerative geometry. *Ann. of Math. (2)*, 155(1):105–129, 2002.
- [2] Alexandre Eremenko and Andrei Gabrielov. An elementary proof of the B. and M. Shapiro conjecture for rational functions. In *Notions of positivity and the geometry of polynomials. Dedicated to the memory of Julius Borcea*, pages 167–178. Basel: Birkhäuser, 2011.
- [3] Lisa R. Goldberg. Catalan numbers and branched coverings by the Riemann sphere. *Adv. Math.*, 85(2):129–144, 1991.

CENTER FOR COMPUTING SCIENCES, INSTITUTE FOR DEFENSE ANALYSES, BOWIE, MD  
*E-mail address:* awfaber@super.org

SMITH COLLEGE, NORTHAMPTON, MA  
*E-mail address:* bthompson@smith.edu